Docket No.:0553-0118.01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:	)	
Shunpei YAMAZAKI	)	
Application No.: 10/694,477	)	Examiner: M. PRENTY
Filed: October 27, 2003	)	Group Art Unit: 2822
For: NONVOLATILE MEMORY AND ELECTRONIC APPARATUS	)	

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on 6/16/06

Shannon Wallace

Date: June 16, 2006

VERIFICATION OF TRANSLATION

Commissioner for Patents
P.O.Box 1450
Alexandria, VA 22313-1450

Sir:

I, Yoko KUNIEDA, C/O Semiconductor Energy Laboratory Co., Ltd. 398, Hase, Atsugi-shi, Kanagawa-ken 243-0036 Japan, a translator, herewith declare:

that I am well acquainted with both the Japanese and English Languages;

that I am the translator of the attached English translation of the Japanese Patent Application No. 09-337710 filed on November 21, 1997; and

that to the best of my knowledge and belief the following is a true and correct English translation of the Japanese Patent Application No. 09-337710 filed on November 21, 1997.

I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Date: this 7th day of June 2006

Yoko Kunieda

Name: Yoko KUNIEDA

[Name of Document] Patent Application

[Reference Number] P003780-01

[Filing Date] November 21, 1997

[Attention] Commissioner, Patent Office

[International Patent Classification] H01L 21/00, H01L 27/10

[Title of Invention] NONVOLATILE MEMORY AND ELECTRONIC APPARATUS

[Number of Claims] 6

[Inventor]

[Address] 398, Hase, Atsugi-shi, Kanagawa-ken

c/o Semiconductor Energy Laboratory Co., Ltd.

[Name] Shunpei Yamazaki

[Applicant]

[Identification Number] 000153878

[Name] Semiconductor Energy Laboratory Co., Ltd.

[Representative] Shunpei YAMAZAKI

[Indication of Handlings]

[Payment Method] Prepayment

[Number of Prepayment Note] 002543

[Payment Amount] 21,000

[List of Attachment]

[Attachment] Specification 1

[Attachment] Drawing 1

[Attachment] Abstract 1

[Document Name] Specification

[Title of Invention]

NONVOLATILE MEMORY AND ELECTRONIC APPARATUS

[Scope of Claims]

[Claim 1]

A nonvolatile memory characterized by including:

a source region, a drain region, and an active region, formed using a single crystal semiconductor;

impurity regions provided locally with respect to the active region; and

an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions.

[Claim 2]

A nonvolatile memory characterized by including:

a source region, a drain region, and an active region, formed using a single crystal semiconductor;

impurity regions provided locally in the active region; and

an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions,

wherein the impurity regions comprise an element selected from group 13 or group 15.

[Claim 3]

A nonvolatile memory characterized by including:

a source region, a drain region, and an active region, formed using a single crystal semiconductor;

impurity regions provided locally in the active region; and

an intrinsic or substantially intrinsic channel forming region interposed

between the impurity regions,

wherein the impurity regions comprise an element selected from group 13 or group 15; and

wherein the impurity regions prevent a depletion layer from expanding from the drain region toward the source region.

[Claim 4]

A nonvolatile memory according to Claims 1 to 3, wherein the impurity regions formed from the source region to the drain region have a stripe shape.

[Claim 5]

A nonvolatile memory according to Claims 1 to 4, wherein a concentration of an element included in the impurity region is in the range of 1×10^{17} to 5×10^{20} atoms/cm^3.

[Claim 6]

An electronic apparatus using the nonvolatile memory according to Claims 1 to 5 as a recording medium.

[Detailed Description of the Invention]

[0001]

[Technical Field to which the Invention pertains]

The present invention relates to the configuration of a nonvolatile memory formed by using a semiconductor. In particular, the invention is effective for a nonvolatile memory in which the channel length is 2 μm or less, or even 0.5 μm or less.

[0002]

[Prior Art]

The IC memories that perform data storage and holding in computers are generally classified into the RAM and the ROM. Examples of the RAM (Random Access Memory) are the DRAM (dynamic RAM) and the SRAM (static RAM). When the power is turned off, data stored in the DRAM or the SRAM are lost.

[0003]

On the other hand, the mask ROM and the PROM (programmable ROM) are known as the ROM, and they have an advantage that even when the power is turned off, data stored there are not lost. The PROM is further classified into the EPROM (Erasable-PROM) in which data erasure is performed by ultraviolet light, the EEPROM (Electrically-EPROM) in which data erasure is performed electrically, the flash memory (flash-EEPROM) in which data erasure is performed en bloc electrically, and the like.

[0004]

To fully utilize their marked advantage of permanent data holding, studies and developments on nonvolatile memories have been made remarkably. Recently, the possibility of using a nonvolatile memory as a substitute memory for a magnetic memory has been discussed.

[0005]

As for such IC memories, it is necessary to increase the storage capacity, at the same time as improving the reliability and performance. That is, as is the case with other types of ICs, miniaturization techniques are always adopted, and development according to the scaling law is advanced.

[0006]

However, basically nonvolatile memories store data using the same principle of operation as field-effect transistors (hereinafter referred to as FETs). Therefore, the short channel effect, which is known as causing serious problems in the FET operation, also causes serious problems in the operation of nonvolatile memories as the miniaturization advances.

[0007]

In particular, the phenomenon called punch-through decreases the source-drain withstand voltage and hence makes the current control with the gate electrode difficult. So, conventionally, there is an example of structure for increasing the punch-through resistance, which is called a SSW-DSA structure (Nikkei Microdevices, pp. 47-48, May issue, 1992).

[0008]

[Problems to be Solved by the Invention]

In the field of the FET, the SSW-DSA structure is a structure that utilizes a technique called a pocket structure, in which an impurity region having the same conductivity type as the substrate is provided in the channel-drain junction portion. In this manner, this structure can suppress the expansion of the drain depletion layer, and prevent the occurrence of punch-through.

[0009]

However, in nonvolatile memories, electron-hole pairs are generated by positively causing impact ionization in the channel-drain junction portion. Therefore, a large amount of holes flow to the substrate side at the same time as injection of electrons into the floating gate.

[0010]

However, in the SSW-DSA structure, a large amount of holes thus generated act in no other way than flow into the substrate terminal. As a result, there may be a problem that a parasitic source-substrate-drain bipolar is formed to cause a kink phenomenon (an abnormal increase in drain current).

[0011]

The present invention has been made in view of the above problems, and an object of the invention is therefore to realize a high-performance memory by effectively preventing or restraining the short channel effect that occurs in miniaturizing nonvolatile memories.

[0012]

[Means for Solving the Problems]

One feature of a structure of the invention disclosed in this specification is characterized by including:

a source region, a drain region, and an active region, formed using a single crystal semiconductor;

stripe-shaped impurity regions provided in the active region; and

an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions.

[0013]

Another feature of a structure of the invention is characterized by including:

a source region, a drain region, and an active region formed using a single crystal semiconductor;

stripe-shaped impurity regions provided in the active region; and

an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions,

wherein the impurity regions comprise an element selected from group 13 or group 15.

[0014]

Another feature of a structure of the invention is characterized by including:

a source region, a drain region, and an active region formed using a single crystal semiconductor;

stripe-shaped impurity regions provided in the active region; and

an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions,

wherein the impurity regions comprise an element selected from group 13 or group 15; and

wherein the impurity regions prevent a depletion layer from expanding from the drain region toward the source region.

[0015]

In the above structures, it is preferable that the impurity regions are formed from the source region to the drain region.

[0016]

Furthermore, in the above structures, it is preferable that a concentration of an element included in the impurity region is in the range of 1×10^{17} to 5×10^{20} atoms/cm^3.

[0017]

Furthermore, it is effective that a recording circuit using a nonvolatile memory with the above structure as a recording medium is formed and incorporated in an electronic apparatus.

[0018]

The main feature of the invention is that impurity regions are formed locally in the active region and the impurity regions prevent a depletion layer from expanding from the drain region toward the source region. In this specification, a region that is enclosed by a source region, a drain region, and field oxide films is called an active region, and the active region is divided into striped impurity regions and channel forming regions.

[0019]

Since the effect of preventing a depletion layer looks like pinning the depletion layer, the inventors define the term "pinning" as meaning "prevention".

[0020]

[Embodiment Mode of the Invention]

Embodiment modes of the invention will be described in detail by the embodiments shown below.

[0021]

[Embodiments]

[Embodiment 1]

This embodiment will be described with reference to FIG. 1. Cross-sectional views and a top view of a nonvolatile memory to which the invention is applied are shown in FIG. 1. In this embodiment, explanation will be made taking an EEPROM having a basic stacked structure for example.

[0022]

In FIG. 1, reference numeral 101 denotes a single crystal silicon (p-type silicon); 102, field oxide films formed by a LOCOS method; 103: a source region formed by adding arsenic (or phosphorus); and 104, a drain region. Although an example structure of an n-type EEPROM is shown here, it is also possible to construct a p-type EEPROM. A p-type EEPROM can be constructed by forming source and drain regions by adding boron to n-type silicon.

[0023]

And, 105 denotes impurity regions (hereinafter referred to as pinning regions) that are the most important feature of the invention. The pinning regions 105 are formed by adding an impurity having the same conductivity type as the silicon substrate 101. In the case of FIG. 1, since p-type silicon is used, the pinning regions 105 are formed by adding an element selected from group 13 (typically boron). Naturally, when n-type silicon is used (when a p-type EEPROM is to be manufactured), pinning regions may be formed by adding an element selected from group 15.

[0024]

The element selected from group 13 or group 15 shifts the energy band of single crystal silicon and thereby forms an energy barrier to carriers (electrons or holes). In this sense, the pinning regions can also be called regions where energy band is shifted. Elements other than the group-13 and group-15 elements can also be used as long as they have such an effect.

[0025]

Here, the elements which shift energy band will be described below with reference to conceptual diagrams shown in FIG. 2. FIG. 2(A) shows an energy band state of single crystal silicon. If an impurity element (an element selected from group 13) that shifts the energy band in such a direction as to obstruct the movement of electrons is added there, the energy band state is changed to a state as shown in FIG. 2(B).

[0026]

At this time, in the impurity-added region, although no change occurs in the energy band gap, the Fermi level (Ef) is moved to the valence band (Ev) side. As a result, the energy is shifted

upward in appearance and hence an energy barrier having higher energy than the undoped regions by ΔE (for electrons) is formed.

[0027]

If an impurity element (an element selected from group 15) that shifts the energy band in such a direction as to obstruct the movement of holes is added to the state of FIG. 2(A), the energy state is changed to a state as shown in FIG. 2(C).

[0028]

In this case, in the impurity-added region, the Fermi level is moved to the conduction band (Ec) side, the energy state is shifted downward in appearance and hence an energy barrier having higher energy than the undoped regions by ΔE (for holes) is formed.

[0029]

As described above, an energy difference corresponding to ΔE occurs between the undoped regions where the impurity is not added and the pinning regions. The height of the energetic (potential) barrier changes depending on the concentration of the added impurity element. In the invention, the impurity element concentration is adjusted in the range of 1×10^{17} to 5×10^{20} atoms/cm^3 (preferably 1×10^{18} to 5×10^{19} atoms/cm^3).

[0030]

Since the formation of the pinning regions 105 is enabled by using microprocessing technology, it is necessary to use an adding means suitable for microprocessing, such as ion implantation or an FIB (focused ion beam) method. Alternatively, when an adding method using a mask is employed, it is desirable to use microprocessing, for instance, forming a mask pattern by using electron beam lithography.

[0031]

Most typically, the pinning regions 105 are formed in such a manner that the pinning regions 105 are approximately parallel with channel forming regions 106 and the pinning regions 105 and the channel forming regions 106 are arranged alternately as shown in FIG. 1(A). That is,

the structure in which a plurality of striped pinning regions are provided in a region (active region) enclosed by the source region 103, the drain region 104, and the field oxide films 102, is preferable.

[0032]

It is effective to form pinning regions at side end portions of the active region (the end portion in which the active region is in contact with the field oxide films). If the pinning regions are formed at the side end portions, leak current that passes through the side end portions can be reduced.

[0033]

It is sufficient to form the pinning regions 105 so that they at least reach the junction portion of the active region and the drain region 104 (the drain junction portion). Since the depletion layer that causes a problem of the punch-through expands from the drain junction portion, the intended effect can be obtained by suppressing the expansion there. That is, the expansion of the depletion layer can be suppressed by forming dot-like or elliptical pinning regions in the active region so that part of them exist in the drain junction portion.

[0034]

Naturally, a more effective pinning effect can be obtained by forming the pinning regions so that they reach both of the source region 103 and the drain region 104, as shown in FIG. 1(A).

[0035]

In addition, it is desirable that the implantation depth of the pinning regions 105 be deeper than the junction depth of the source/drain region. Therefore, the implantation depth of 0.1 to 0.5 μm (preferably 0.2 to 0.3 μm) is required.

[0036]

Now, the channel length and the channel width will be defined with reference to FIG. 3. In FIG. 3, the distance between a source region 301 and a drain region 302 (which corresponds to the length of an active region 303) is defined as a channel length (L). The invention is effective in a case where this length is 2 μm or less, typically 0.05 to 0.5 μm, and preferably 0.1 to 0.3 μm.

The direction along this channel length is called a channel length direction.

[0037]

The width of an arbitrary pinning region 304 is called a pinning width (v_j). The pinning width may be set to 1 μm or less, typically 0.01 to 0.2 μm, and preferably 0.05 to 0.1 μm. When the sum of the widths of all pinning regions existing in the active region 303 is set to be an effective pinning width (V), it is defined as the following formula.

[0038]

[Formula 1]

[0039]

To obtain the pinning effect, it is necessary to form at least one pinning region in the active region 303. That is, a condition of j=1 or more needs to be satisfied. In the case where pinning regions are provided at the side end portions (portions in contact with field oxide films) of the active region 303, at least a condition of j=2 or more needs to be satisfied.

[0040]

The width of a channel forming region 305 is set to be a channel width (w_i). Although the channel width can respond to any case, for memories in which large current is not required, it may be set to 1 μm or less, typically 0.05 to 0.5 μm, and preferably 0.1 to 0.3 μm.

[0041]

In addition, when the sum of all channel widths (w_i) is set to be an effective channel width (W), it is defined as the following formula.

[0042]

[Formula 2]

[0043]

In the case where the pinning regions are formed only at the side end portions of the active region 303, i=1 is satisfied. In order to obtain the pinning effect effectively, the pinning regions are preferably formed also in regions other than the side end portions of the active region 303. In

such a case, i becomes 2 or more.

[0044]

In addition, a total channel width (W_{total}) that is the sum of the sum of the all pinning regions (the effective pinning width) and the sum of the all channel forming regions (the effective channel width) is defined as the following formula.

[0045]

[Formula 3]

[0046]

The total channel width (W_{total}) corresponds to the width of the active region 303 (the length of the active region in the direction perpendicular to the channel length direction). The direction along this total channel width is called a channel width direction.

[0047]

Since the invention is intended to be applied to nonvolatile memories having an extremely short channel length as mentioned above, the pinning regions and the channel forming regions need to be formed so as to have extremely small dimensions.

[0048]

In FIG. 1, it is preferable that the impurity element that has been added to the pinning region 105 is activated by furnace annealing, laser annealing, lamp annealing or the like. This activation step may be performed at the same time as annealing treatment in a later step such as formation of a gate insulating film, or independently of such annealing treatment.

[0049]

A feature of the invention is that pinning regions are formed locally (in striped form) in a region which functions as a channel forming region in a conventional nonvolatile memory. Therefore, the other structures of the conventional nonvolatile memory can be employed as they are.

[0050]

That is, a tunnel oxide film 107 is provided on the single crystal silicon on which the

source region 103, the drain region 104, and the pinning regions 105 are provided. The tunnel oxide film, which is formed by a thermal oxidation step, is required to have high film quality. In this embodiment, the thickness of the tunnel oxide film 107 is set to 11 nm. It goes without saying that the thickness of the tunnel oxide film is not limited to this value.

[0051]

In this embodiment, the above-described pinning regions 105 may be formed after the tunnel oxide film 107 is formed.

[0052]

A floating gate 108 formed of a first polysilicon layer is provided on the tunnel oxide film 107. As shown in FIG. 1(C), it is important to make a structure in which the end portions of the floating gate 108 overlap the junction portions of the pinning portions 105 and the drain region 104.

[0053]

A large amount of hot electrons are generated by concentrated electric fields at the junction portions. Therefore, hot electrons can be injected at high efficiency by making the floating gate overlapping the portions.

[0054]

After the floating gate 108 is provided in this manner, a first interlayer film 109, a control gate 110 formed of a second polysilicon layer, a second interlayer film 111, and a bit line 112 are provided.

[0055]

Naturally, a conductive layer such as a metal film can be used instead of the polysilicon layer. In addition, it is also effective to use, as the interlayer film, a laminated film as expressed by $SiO_2/SIN/SiO_2$ (commonly called an ONO film).

[0056]

The two-layer polysilicon EEPROM of this embodiment is expressed by a circuit diagram shown in FIG. 1(D). In FIG. 1(D), Vd denotes a drain voltage, Vs denotes a source voltage, C.G.

denotes a control gate voltage, and F.G. denotes a potential of the floating gate.

[0057]

In the EEPROM of this embodiment, the following voltages are applied at the time of data writing and erasure.

[0058]

[Table 1]

[0059]

Naturally, the operation voltages are not limited to Table 1. Furthermore, the invention is not limited to the structure of this embodiment and can be applied to any kinds of EEPROMs in which data is erased electrically.

[0060]

(Functions and advantages of the invention)

First, a first advantage of the invention will be described below. In FIG. 1, the pinning regions 105 that are formed locally in the active region function as stoppers with respect to the depletion layer that expands from the drain side and effectively suppress expansion of the depletion layer. Therefore, the punch-through phenomenon that is caused by expansion of the depletion layer can be prevented. Furthermore, since an increase of depletion layer charge due to expansion of the depletion layer is suppressed, a reduction in threshold voltage can be also prevented.

[0061]

Next, a second advantage will be described. In the invention, the narrow channel effect can be enhanced intentionally by the pinning regions. The narrow channel effect, which is a phenomenon observed when the channel width is extremely narrow, causes an increase in threshold voltage (refer to Submicron Devices I, Mitsumasa Koyanagi et al., pp. 88-138, Maruzen Co., Ltd., 1987).

[0062]

FIG. 4 shows an energy state (potential state) of the active region when the pinning TFT of

this embodiment operates. In FIG. 4, regions denoted by 401 and 402 correspond to the energy state of pinning regions 105, and a region denoted by 403 corresponds to the energy state of a channel forming region 106.

[0063]

As seen from FIG. 4, the pinning regions 105 form high-energy barriers, and the channel forming regions 106 form low-energy regions. Therefore, carriers move through the channel forming regions 106 with priority where the energy state is low.

[0064]

In this manner, high-energy barriers are formed in the pinning regions 105 and the threshold voltage increases there. As a result, a threshold voltage that is observed as a whole also increases. This narrow channel effect becomes more remarkable as the effective channel width decreases.

[0065]

As described above, in the invention, by freely designing the concentration of an impurity that is added to the pinning regions 105 and the effective channel width, the degree of the narrow channel effect is controlled and the threshold voltage can be adjusted. That is, by controlling the pinning effect, a threshold voltage decrease due to the short channel effect and a threshold voltage increase due to the narrow channel effect are balanced, and the threshold voltage can be adjusted to a desired value.

[0066]

In addition, since a group-13 element is added to the pinning regions in an n-type device and a group-15 element is added in a p-type device, the threshold voltage is shifted in that portion in a direction of increasing (in the positive direction in an n-type device and in the negative direction in a p-type device). That is, since the threshold voltage increases locally, the threshold voltage of the entire device increases accordingly. Therefore, in order to adjust the threshold voltage to a desired value, it is important to set the concentration of an impurity that is added to the pinning

regions at a proper value.

[0067]

Incidentally, in nonvolatile memories, discrimination between "0" and "1" is made by changing the threshold voltage by injecting charge (mainly electrons) into the floating gate and detecting whether current flows through the bit line when a certain voltage is applied. Therefore, if the threshold voltage becomes extremely small by the short channel effect, it becomes necessary to discriminate between "0" and "1" by applying a very small voltage. That is, the device is easily influenced by noise or the like and the possibility of an erroneous operation increases.

[0068]

However, in the invention, since the threshold voltage can be controlled to have a desired value by suppressing a threshold voltage reduction, the ability of discriminating between "0" and "1" is increased. Therefore, a nonvolatile memory with very high reliability can be realized.

[0069]

Next, a third advantage will be described. The nonvolatile memory according to the invention has an advantage that the channel forming regions 106 are structured as substantially intrinsic regions, and majority carriers (electrons in an n-type device and holes in a p-type device) move through these regions.

[0070]

Here, "substantially intrinsic region" basically means an undoped single crystal semiconductor region, and includes a region where conductivity type compensation is made intentionally by adding an impurity element of the opposite conductivity type, and a region having one conductivity type in a range where the threshold voltage can be controlled.

[0071]

For example, it can be said that a silicon wafer in which a dopant concentration is 5×10^{16} atoms/cm^3 or less (preferably 5×10^{15} atoms/cm^3 or less) and a concentration of contained carbon, nitrogen, and oxygen is 2×10^{18} atoms/cm^3 or less (preferably 5×10^{17} atoms/cm^3 or less) is

substantially intrinsic. In this sense, silicon wafers commonly used are substantially intrinsic unless an impurity is added intentionally in a process.

[0072]

In the case where a region in which carriers move is substantially intrinsic, a mobility reduction due to the impurity scattering is extremely small and hence high carrier mobility is obtained. That is, the carrier mobility is dominated by the influence of the lattice scattering, and a state that is very close to the ideal is obtained.

[0073]

In the case where the linear pinning regions are formed so as to reach both of the source region and the drain region, as shown in FIG. 1(A), an advantage that movement paths of majority carriers are defined by the pinning regions is obtained.

[0074]

As described above, energy state of the channel forming region interposed between pinning regions is as shown in FIG. 4. In the structure of FIG. 1(A), a plurality of slits having the energy state of FIG. 4 are considered to be arranged.

[0075]

FIG. 5 schematically illustrates such a state. In FIG. 5, 501 denotes pinning regions, 502 denotes channel forming regions, and 503 denotes majority carriers (electrons or holes). As shown in FIG. 5, carriers 503 cannot go over the pinning regions 501 and hence move through the channel forming regions 502 with priority. That is, the movement paths of the majority carriers are defined by the pinning regions.

[0076]

By defining the movement paths of majority carriers, scattering due to self-collision of carriers are decreased, which greatly contributes to mobility increase. Furthermore, since only a very small amount of impurity elements exist in the substantially intrinsic channel forming regions, the velocity overshoot effect occurs, which is a phenomenon that the electron mobility becomes

higher than usual even at room temperature (refer to K. Ohuchi et al., Jpn. J. Appl. Phys. 35, pp. 960, 1996). Therefore, the mobility becomes extremely high.

[0077]

High carrier mobility that is obtained as described above is effective in shortening the charge write time and the charge read time, thereby increasing the memory performance. In addition, high carrier mobility means having high energy, and hence the charge writing efficiency is greatly increased by channel hot electron injection (CHE injection).

[0078]

Next, a fourth advantage of the invention will be described. In the case where the structure of the invention is employed, an electric field is concentrated to a large extent at the junction portions (typically, a p^+/n^{++} junction or an n^+/p^{++} junction is formed) of the pinning regions and the drain region. Therefore, there occur a large amount of electrons that are accelerated and have high energy or electrons that are generated by impact ionization (collectively called hot electrons).

[0079]

That is, charge injection into the floating gate is performed very efficiently and hence the data write time can be shortened; specifically, it can be reduced to 1/10 to 1/100 of that of conventional cases. Therefore, by utilizing the invention, the data write time of a 256 Mbit stacked flash memory, which will be realized in the future, can be made 10 ns/byte or less, preferably 0.01 to 1 ns/byte.

[0080]

Since this means realization of a data write operation faster than that of magnetic memories, it becomes possible to replace all kinds of current hard disks structured by magnetic memories with flash memories. That is, very small chips will have functions equivalent to conventional magnetic memories, and it is expected that the miniaturization and the price reduction of devices will be accelerated.

[0081]

Furthermore, the invention is also effective in decreasing the voltage necessary for data writing (the write voltage). That is, since hot electron injection is accelerated due to concentrated electric fields at the pinning/drain junction portions, charge of the same amount as in conventional cases can be injected by a write voltage that is 1/2 to 1/10 of conventional values.

[0082]

Therefore, while stacked flash memories require a write voltage of about 10 V under present circumstances, the invention can realize a write voltage of 5 V or less, preferably about 1 to 3 V.

[0083]

In this manner, the hot electron injection efficiency at the drain junction portions can be increased by providing the pinning regions. This is effective in reducing the power consumption and increasing the degree of freedom in circuit designing.

[0084]

Next, a fifth advantage will be described. The fact that the pinning regions of the invention have the functions of preventing the short channel effect and controlling the threshold voltage has been described above. Besides these, the pinning regions of the invention have a very important role in preventing a parasitic bipolar from being rendered conductive due to impact ionization (collisional ionization).

[0085]

Conventionally, electrons of electron-hole pairs generated by impact ionization are injected into the floating gate. On the other hand, holes flow into the substrate and cause a substrate current, which renders a parasitic bipolar conductive.

[0086]

However, in the invention, holes generated by impact ionization immediately move into the pinning regions, and extracted to the source region through the inside of the pinning regions.

Therefore, a parasitic bipolar is not rendered conductive and hence the source-drain withstand voltage is not reduced.

[0087]

It goes without saying that this effect is particularly remarkable when the pinning regions are formed so as to reach both of the source and drain regions. In addition, holes can be extracted more efficiently if the pinning regions are in contact with a pickup electrode in the source region.

[0088]

[Embodiment 2]

The two-layer polysilicon EEPROMs described in the first embodiment can be classified into a byte erasure type (data erasure is performed on a unit memory element basis) and a flash type (data of collective memory elements are erased en bloc).

[0089]

The flash EEPROM is also called a flash memory. The invention can be applied to either of the two types of EEPROMs.

[0090]

There are various data erasing methods such as a source erasure type, a source/gate erasure type, and a substrate erasure type. The invention can be applied to any of these methods.

[0091]

[Embodiment 3]

Examples of a two-layer polysilicon EEPROM are described in the first and second embodiments, and this embodiment will explain a case where the invention is applied to a three-layer polysilicon EEPROM, with reference to FIG. 6.

[0092]

Since an EEPROM of this embodiment has the same basic structure as the two-layer polysilicon EEPROM described in the first embodiment, the reference numerals used in FIG. 1 are also used in this embodiment. That is, for the parts shown in FIG. 6 with the same reference

numerals as in FIG. 1, reference is made to the descriptions relating to FIG. 1. In this embodiment, only different parts will be given new reference numerals and described below.

[0093]

FIG. 6(A) is different from FIG. 1(A) in that an erasing gate 601 is provided. That is, a first polysilicon layer constitutes the erasing gate 601, a second polysilicon layer constitutes the floating gate 108 and a third polysilicon layer constitutes the control gate 110.

[0094]

In the EEPROM with a structure of the first embodiment, data erasure is performed by extracting, to the substrate side (the source region or the bulk substrate), electrons that have been injected into the floating gate 108. In contrast, in the EEPROM with a structure of this embodiment, data erasure is performed by extracting, to the erasing gate 601, electrons that have been injected into the floating gate 108.

[0095]

Therefore, in FIG. 1(B), an insulating film 602 for insulating the erasing gate 601 and the floating gate 108 from each other should be as thin as possible (preferably 8 to 12 nm) so as to allow a flow of tunnel current (Fowler-Nordheim current) and should be of such high quality as to be highly durable.

[0096]

The EEPROM of this embodiment can be manufactured basically by the same process as the structure shown in the first embodiment, with an exception that a step of forming the erasing gate 601 and the erasing gate insulating film 602 after formation of the pinning regions is added.

[0097]

An EEPROM having an erasing gate like the one of this embodiment is expressed by a circuit diagram shown in FIG. 6(D), in which Vd denotes a drain voltage, Vs denotes a source voltage, E.G. denotes an erase gate voltage, C.G. denotes a control gate voltage, and F.G. denotes a floating gate potential.

[0098]

In the EEPROM of this embodiment, the following voltages are applied at the time of data writing and erasure.

[0099]

[Table 2]

[0100]

Naturally, the operation voltages are not limited to Table 2. Furthermore, the invention is not limited to the structure of this embodiment and can be applied to any kinds of EEPROMs having an erasing gate structure.

[0101]

[Embodiment 4]

The nonvolatile memories described in the first to third embodiments utilize hot electron injection for data writing, and Fowler-Nordheim current for data erasure. This type of operation mode is called a stack mode.

[0102]

The invention can also be applied to nonvolatile memories using Fowler-Nordheim current for data writing. This type of operation mode is classified into a NAND type, an AND type, a DINOR type and the like.

[0103]

In particular, when manufacturing a large-capacity memory of 256 Mbits or more, to improve the reliability (to elongate the life by decreasing deterioration) it is preferable to write data by using Fowler-Nordheim current.

[0104]

[Embodiment 5]

As the two-layer polysilicon structure described in Embodiment 1, the description is made taking an EEPROM in which data erasure is performed electrically for example. On the other

hand, the nonvolatile memory in which electrons that have been injected into the floating gate are extracted to the source or the substrate by exciting those by ultraviolet irradiation or heating is called an EPROM. The invention can also be applied to the EPROM.

[0105]

Among various kinds of EPROMs is a nonvolatile memory not using a floating gate in which a two-layer gate insulating film is provided between a control gate and a silicon substrate and hot electrons are captured by interface states of the two-layer gate insulating film. For example, a nonvolatile memory of a type in which hot carriers are captured at the interface between a silicon oxide film and a silicon nitride film is called an NMOS nonvolatile memory.

[0106]

Furthermore, there are nonvolatile memories in which hot carriers are captured by metal clusters, silicon clusters, or the like that are formed intentionally at an interface between insulating films.

[0107]

The invention can be applied to all kinds of EPROMs as described above.

[0108]

[Embodiment 6]

Since the invention is applicable to all kinds of conventional nonvolatile memories, the circuit configuration of the invention can be applied to all the known circuit configurations. This embodiment will describe a case where the invention is applied to flash memories that are designed according to the NAND and NOR type architectures, respectively.

[0109]

First, a description will be made of the configuration of a NAND-type memory circuit shown in FIGS. 7(A) and 7(B). It is to be noted that, in FIG. 7, two circuits each composed of eight memory transistors and two selection transistors are shown, but the description will be made of one of them.

[0110]

In FIG. 7(A), reference numerals 701 and 702 denote selection transistors, and they have selection lines S1 and S2 denoted by 703 and 704 as gate electrodes, respectively. The selection transistor 701 connects a bit line 705 denoted by B1 (or B2) to eight memory transistors 706 to 713.

[0111]

Although a case where eight memory transistors are connected to each other in series is shown in this embodiment, the number of memory transistors is not limited.

[0112]

The selection transistor 702 is connected to a final-stage memory transistor 713. The other terminal of the selection transistor 702 is grounded. Even if it is connected to a power supply line, instead of being grounded, operation is still possible, of course.

[0113]

The memory transistors 706 to 713 use word lines 714 to 721 (denoted by W1 to W8) as control gates, respectively.

[0114]

When the NAND-type memory circuit of FIG. 7(A) is shown schematically as a circuit pattern, it is as shown in FIG. 7(B). In each memory transistor, hatched regions are floating gates that are provided under the control gates 714 to 721.

[0115]

Next, a description will be made of the configuration of a NOR-type memory circuit shown in FIGS. 8(A) and 8(B). In FIG. 8, two circuits each composed of 4 memory transistors are shown, but the description will be made of one of them.

[0116]

In FIG. 8(A), four memory transistors 802 to 805 are individually connected to a bit line 801 denoted by B1. Those terminals (source regions) of the respective memory transistors 802 to 805 which are not connected to the bit line 801 are connected to a ground line 806.

[0117]

The memory transistors 802 to 805 use word lines 807 to 810 denoted by W1 to W4 as control gates, respectively.

[0118]

When the NOR-type memory circuit of FIG. 8(A) is shown schematically as a circuit pattern, it is as shown in FIG. 8(B). In the respective memory transistors, hatched regions are floating gates that are provided under the control gates 807 to 810.

[0119]

Although NAND-type circuits as shown in FIG. 7 have disadvantages that the order of writing is fixed and the read access time is long, they have an advantage that the integration density can greatly be increased.

[0120]

The NOR-type circuit shown in FIG. 8 is effective in injecting electrons precisely into the floating gates and reading out charge amounts precisely. This is the feature of the NOR type architecture in which individual memory transistors are directly connected to a source-drain bus line.

[0121]

Although this embodiment describes the nonvolatile memories that use electrodes with the two-layer structure (polysilicon or the like), it can be implemented in nonvolatile memories having electrodes with the three-layer structure (the structure including the erasing gate) as described in the third embodiment.

[0122]

[Embodiment 7]

In this embodiment, a case where a nonvolatile memory according to the invention is applied to a microprocessor that is integrated on one chip, such as a RISC processor or an ASIC processor, will be described.

[0123]

In FIG. 9, an example of a microprocessor is shown. The microprocessor is typically composed of a CPU core 11, a flash memory 12 (or a RAM), a clock controller 13, a cache memory 14, a cache controller 15, a serial interface 16, an I/O port 17, and the like.

[0124]

The microprocessor shown in FIG. 9 is a simplified example, naturally, and a variety of circuit designs are employed in actual microprocessors in accordance with their uses.

[0125]

In the microprocessor shown in FIG. 9, the CPU core 11, the clock controller 13, the cache controller 15, the serial interface 16, and the I/O port 17 are constituted of CMOS circuits 18. Pinning regions 19 disclosed in the invention are provided in the CMOS circuits 18.

[0126]

In this manner, the invention can be applied to MOSFETs, as well as nonvolatile memories. As for the details, application has been made already as Japanese Patent Laid-Open No. Hei. 8-232553.

[0127]

In addition, the nonvolatile memories according to the invention are used in the flash memory 14 and a memory circuit 20 is structured. Every memory cell that constitutes the memory circuit 20 is provided with pinning regions 21. It is possible to use nonvolatile memories according to the invention in the cache memory 12.

[0128]

As described above, FIG. 9 is an example of a case where the pinning technology disclosed in the invention is utilized in all of the memory sections and other logic sections.

[0129]

Furthermore, in some cases, a configuration shown in FIG. 10 may be employed. FIG. 10 shows an example of a case where the logic sections excluding the memory sections are structured

by ordinary CMOS circuits 22. In this case, a structure in which pinning regions are not provided in the logic sections only may be employed.

[0130]

In this manner, it is possible to provide pinning regions in the necessary part, at the circuit designing stage, and a party who practices the invention may determine arbitrarily whether to use pinning regions in the entire circuit or only part of it. In the case where the invention is applied to a hybrid IC in which various functions are combined, such a high degree of freedom in circuit design is very effective.

[0131]

[Embodiment 8]

A semiconductor circuit (memory circuit) structured by the nonvolatile memory according to the invention can be incorporated, as a recording medium for data storage and readout, in electronic apparatuses of every field. In this embodiment, examples of those electronic apparatuses are shown in FIG. 11.

[0132]

As examples of an electronic apparatus which can use a nonvolatile memory of this invention, a video camera, an electronic still camera, a projector, a head-mounted display, a car navigation apparatus, a personal computer, portable information terminals (a mobile computer, a cellular telephone, a PHS, and the like) are given.

[0133]

FIG. 11(A) shows a cellular telephone, which is composed of a main body 2001, a voice output portion 2002, a voice input portion 2003, a display device 2004, manipulation switches 2005, and an antenna 2006. This invention is incorporated in a built-in LSI board, and used to add an address function for recording telephone numbers, or the like.

[0134]

FIG. 11(B) shows a video camera, which is composed of a main body 2101, a display

device 2102, a sound input portion 2103, manipulation switches 2104, a battery 2105, and an image receiving portion 2106. This invention is incorporated in a built-in LSI board, and used for a function of image data storage, or the like.

[0135]

FIG. 11(C) shows a mobile computer, which is composed of a main body 2201, a camera section 2202, an image receiving portion 2203, a manipulation switch 2204, and a display device 2205. This invention is incorporated in a built-in LSI board, and used for storage of processed data and image data.

[0136]

FIG. 11(D) shows a head-mounted display, which is composed of a main body 2301, display devices 2302, and a band portion 2303. This invention is connected to the display devices 2302, as an image signal correction circuit.

[0137]

FIG. 11(E) shows a rear type projector, which is composed of a main body 2401, a light source 2402, a display device 2403, a polarizing beam splitter 2404, reflectors 2405 and 2406, and a screen 2407. This invention can be used as a storage circuit for storing data to be given to a γ-correction circuit.

[0138]

FIG. 11(F) shows a front type projector, which is composed of a main body 2501, a light source 2502, a display device 2503, an optical system 2504, and a screen 2505. This invention can be used as a storage circuit for storing data to be given to a γ-correction circuit.

[0139]

As described above, the application range of the invention is extremely wide and the invention can be applied to electronic apparatuses of every field. In addition to the above examples, the invention can be used as a storage medium that is indispensable in various control circuits and information processing circuits.

[0140]

[Effects of the Invention]

By using the invention, influences of the miniaturization effects as typified by the short channel effect can be minimized, and further miniaturization of nonvolatile memories can be advanced.

[0141]

In addition, nonvolatile memories realizing small-area and large-capacity can be realized, while securing their high reliability.

[Brief Description of the Drawings]

[FIG. 1] Diagrams showing the structure of a nonvolatile memory according to the present invention

[FIG. 2] Diagrams illustrating changes of an energy band

[FIG. 3] A diagram illustrating definitions of a channel length and a channel width

[FIG. 4] A diagram illustrating the energy state in an active region

[FIG. 5] A diagram illustrating the energy state in an active region

[FIG. 6] Diagrams showing the structure of a nonvolatile memory according to the invention

[FIG. 7] Diagrams showing a circuit using a nonvolatile memory according to the invention

[FIG. 8] Diagrams showing a circuit using a nonvolatile memory according to the invention

[FIG. 9] A diagram showing a semiconductor circuit using a nonvolatile memory according to the invention

[FIG. 10] A diagram showing a semiconductor circuit using a nonvolatile memory according to the invention

[FIG. 11] Diagrams showing electronic apparatuses each using a nonvolatile

memory of the invention

[Document Name] Abstract

[Summary]

[Problem] A high-performance memory is realized by efficiently restraining a short channel effect which occurs as a nonvolatile memory is miniaturized.

[Solving Means] In the nonvolatile memory, pinning regions 105 are provided locally in an active region that is enclosed by field oxide films 102, a source region 103 and a drain region 104. In the present invention, a depletion layer expanding from the drain side toward the source side is restrained by the pinning regions 105, so that a punch-through phenomenon accompanied by the short channel effect is prevented.

[Selected Drawing] FIG. 1

Reference No. P003780-01
Document where chemical formulas etc. are descibed
Name of Document Specification

[Formula 1]

$$V = \sum_{j=1}^{n} v_j$$

[Formula 2]

$$W = \sum_{i=1}^{m} w_i$$

[Formula 3]

$$W_{total} = V + W$$

Reference No. P003780-01
Document where chemical formulas etc. are descibed
Name of Document Specification

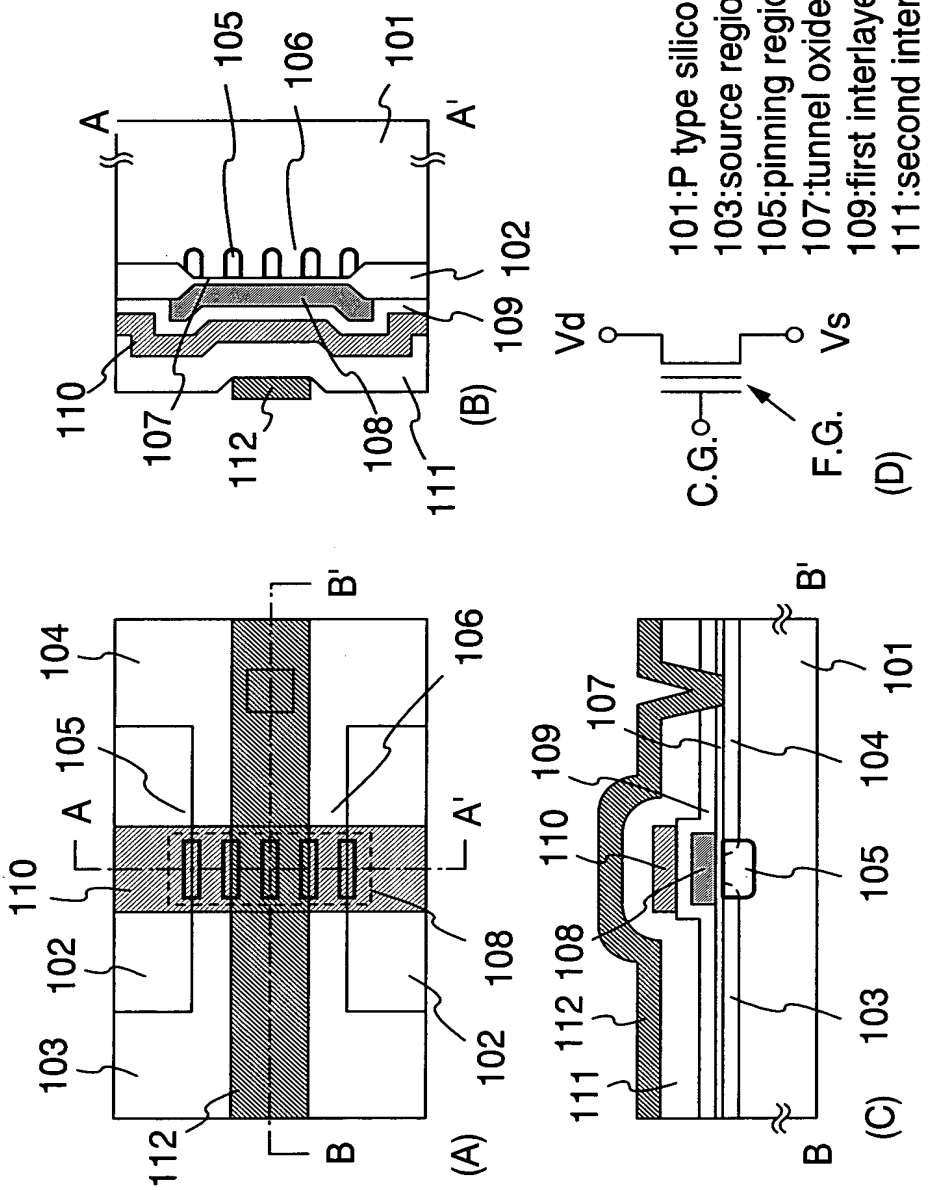
[Table 1]

mode	Vd	V_{CG}	Vs	mechanism
at writing	6	12	0	hot electron injection
at erasing	-	0	12	F-N tunnel erasure
at readout	~1	5	0	-

[Table 2]

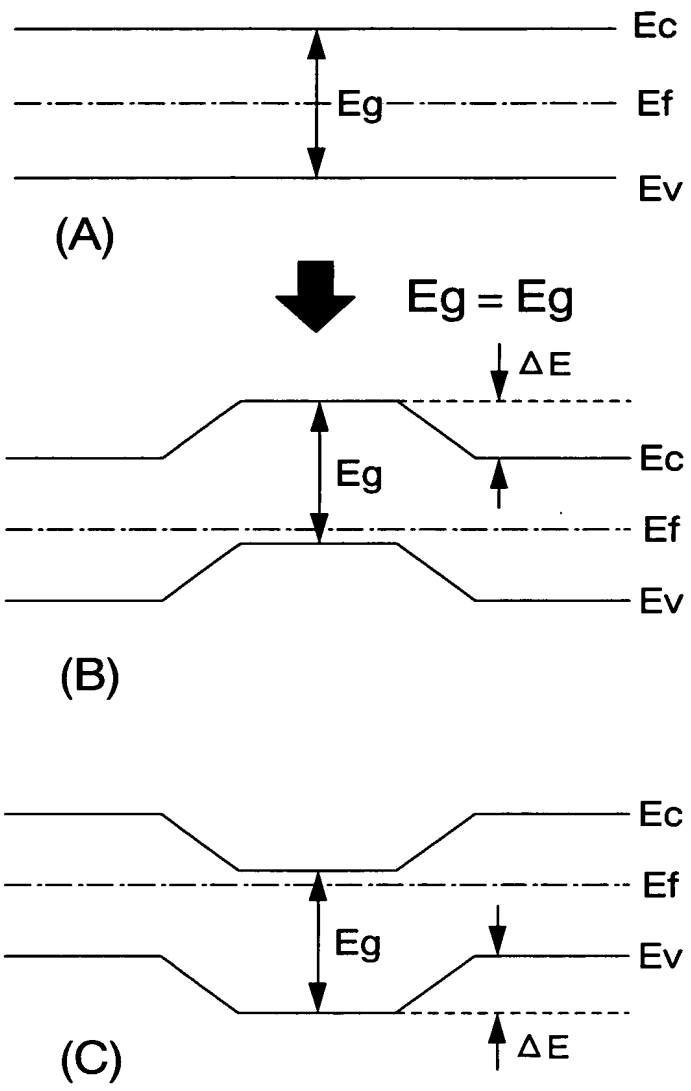
mode	Vd	Vs	V_{EG}	V_{CG}	mechanism
at writing	8	0	3	12	hot electron injection for floating gate
at erasing	-	0	20	0	F-N tunnel erasure from floating gate
at readout	1	0	0	5	-

[Reference No.] P003780-01
[Name of Document] Drawing
[Fig. 1]

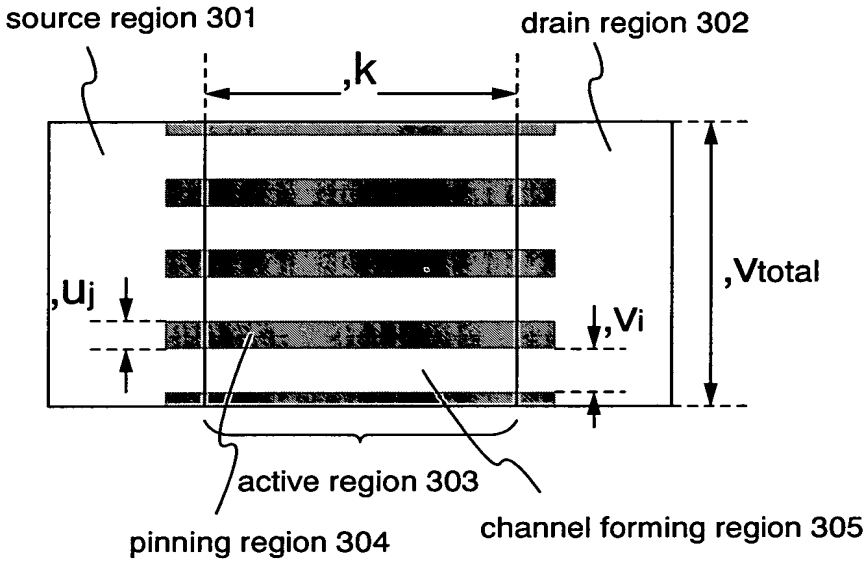


[Refernce No.] P003870-01

[Fig. 2]

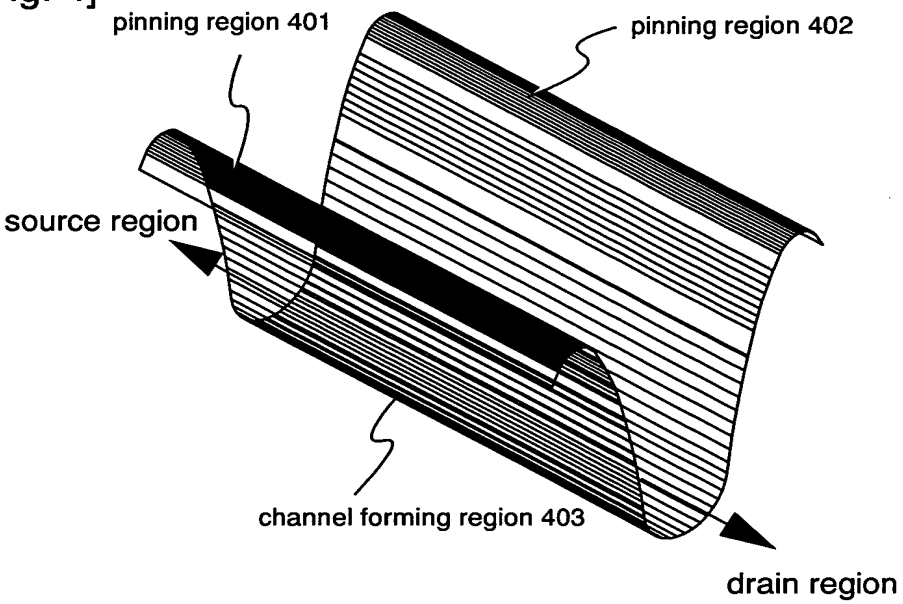


[Reference No.] P003780-01
[Fig. 3]

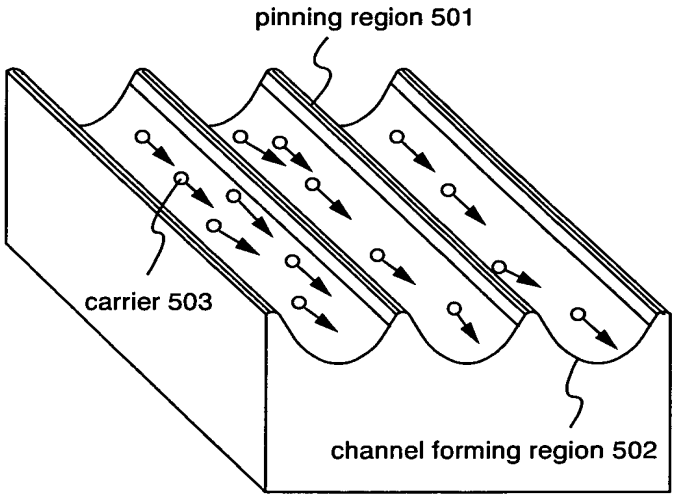


[Reference No.] P003780-01

[Fig. 4]

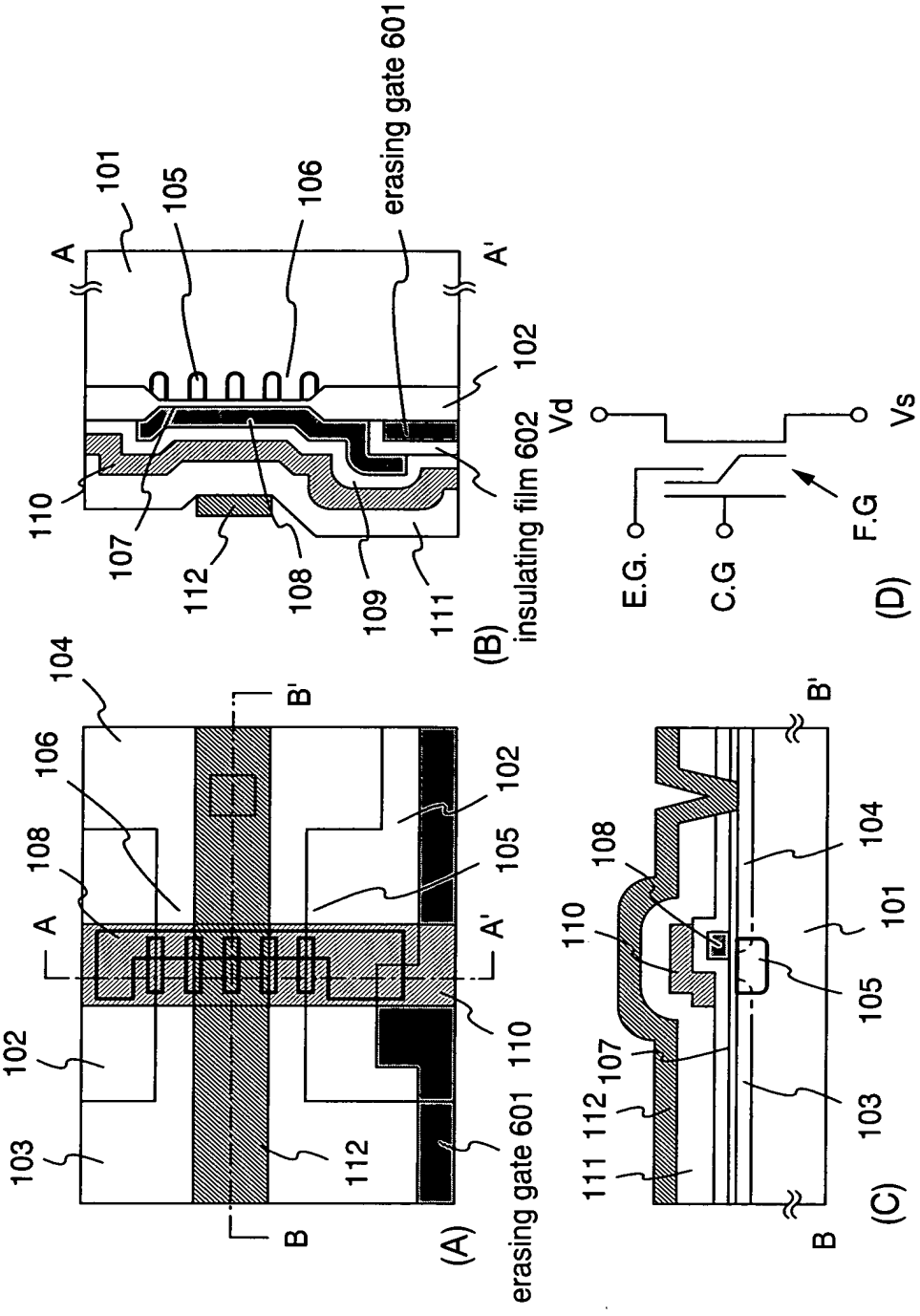


[Fig. 5]



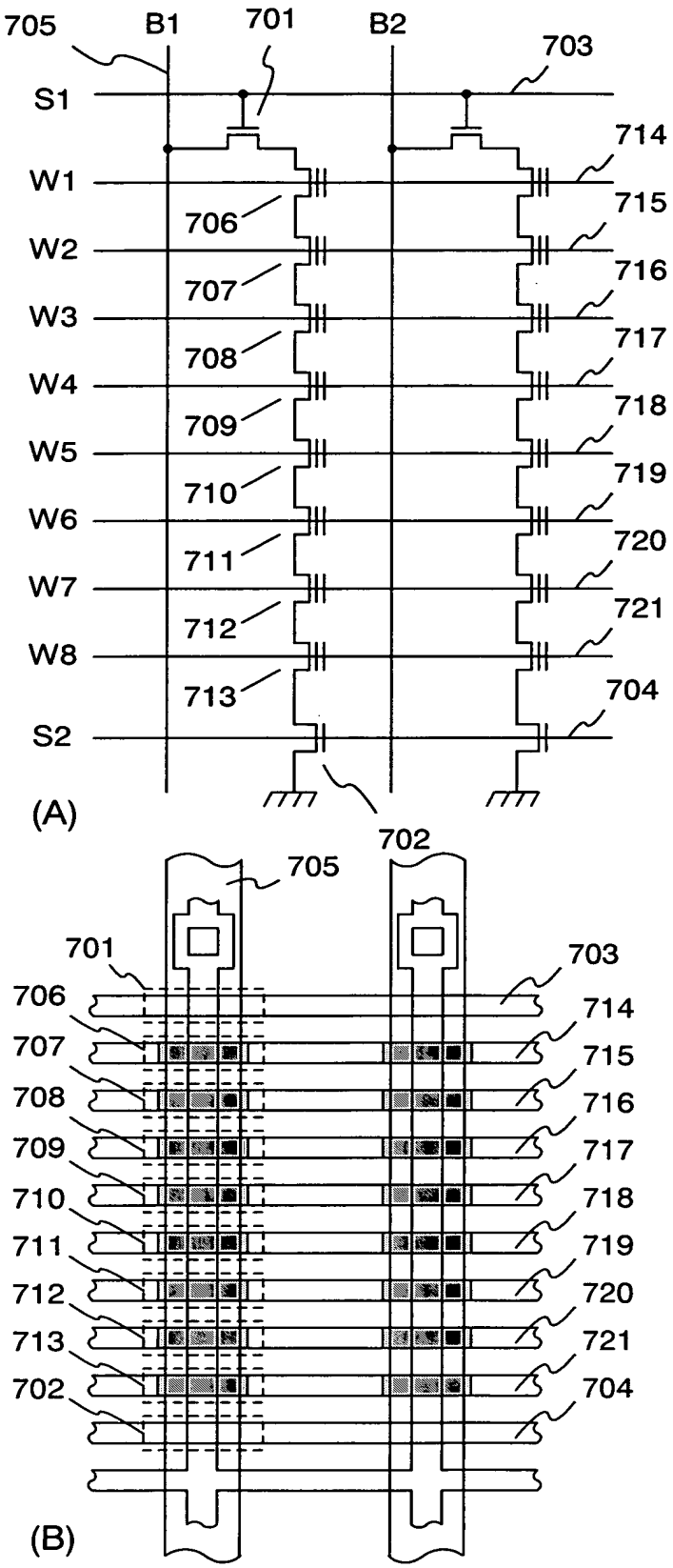
[Reference No.] P003780-01

[Fig. 6]

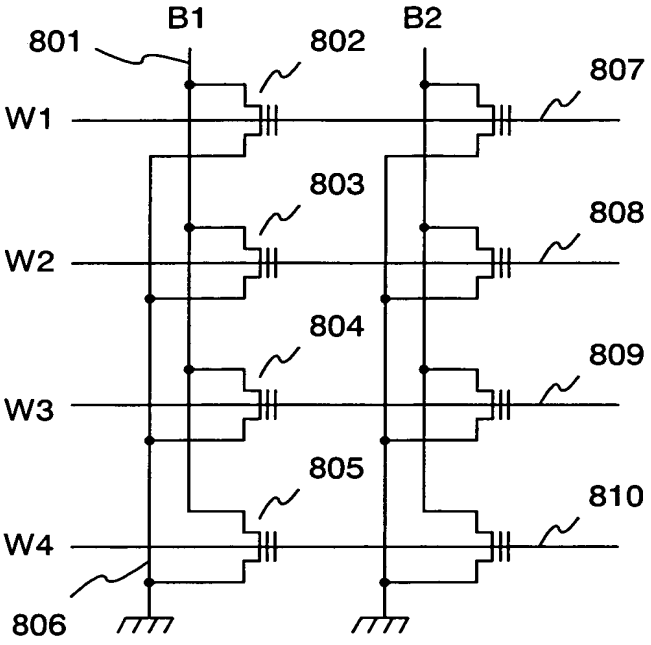


[Reference No.] P003780-01

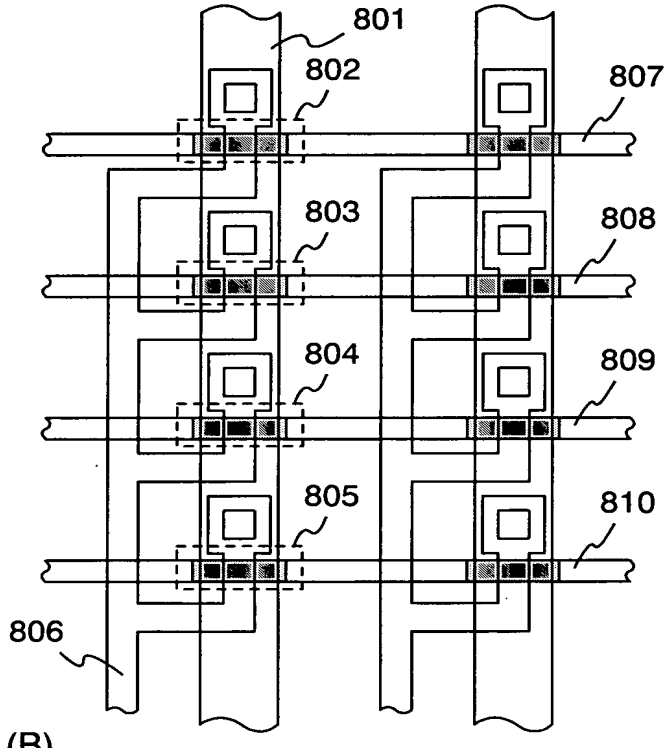
[Fig. 7]



[Reference No.] P003780-01
[Fig. 8]

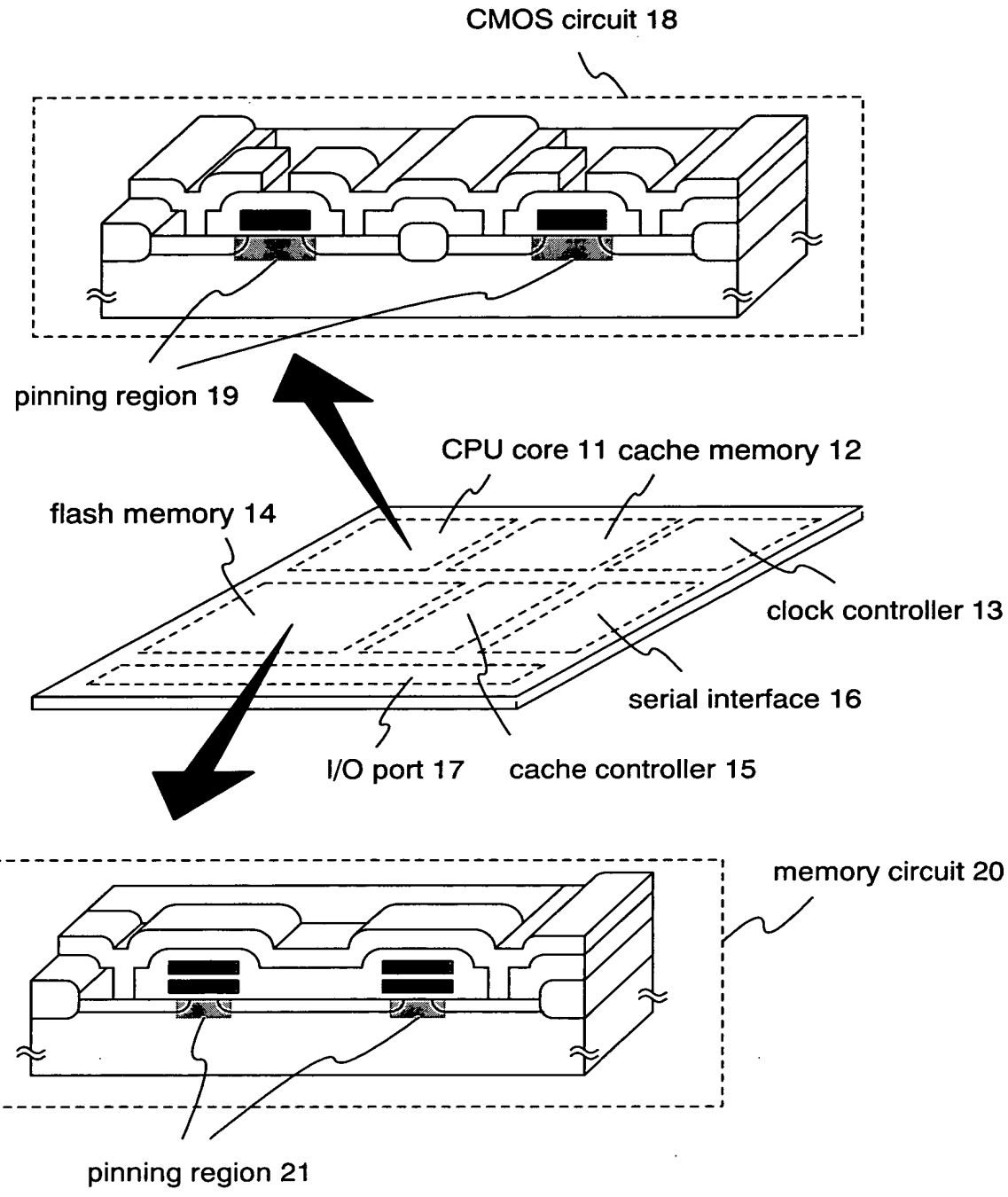


(A)

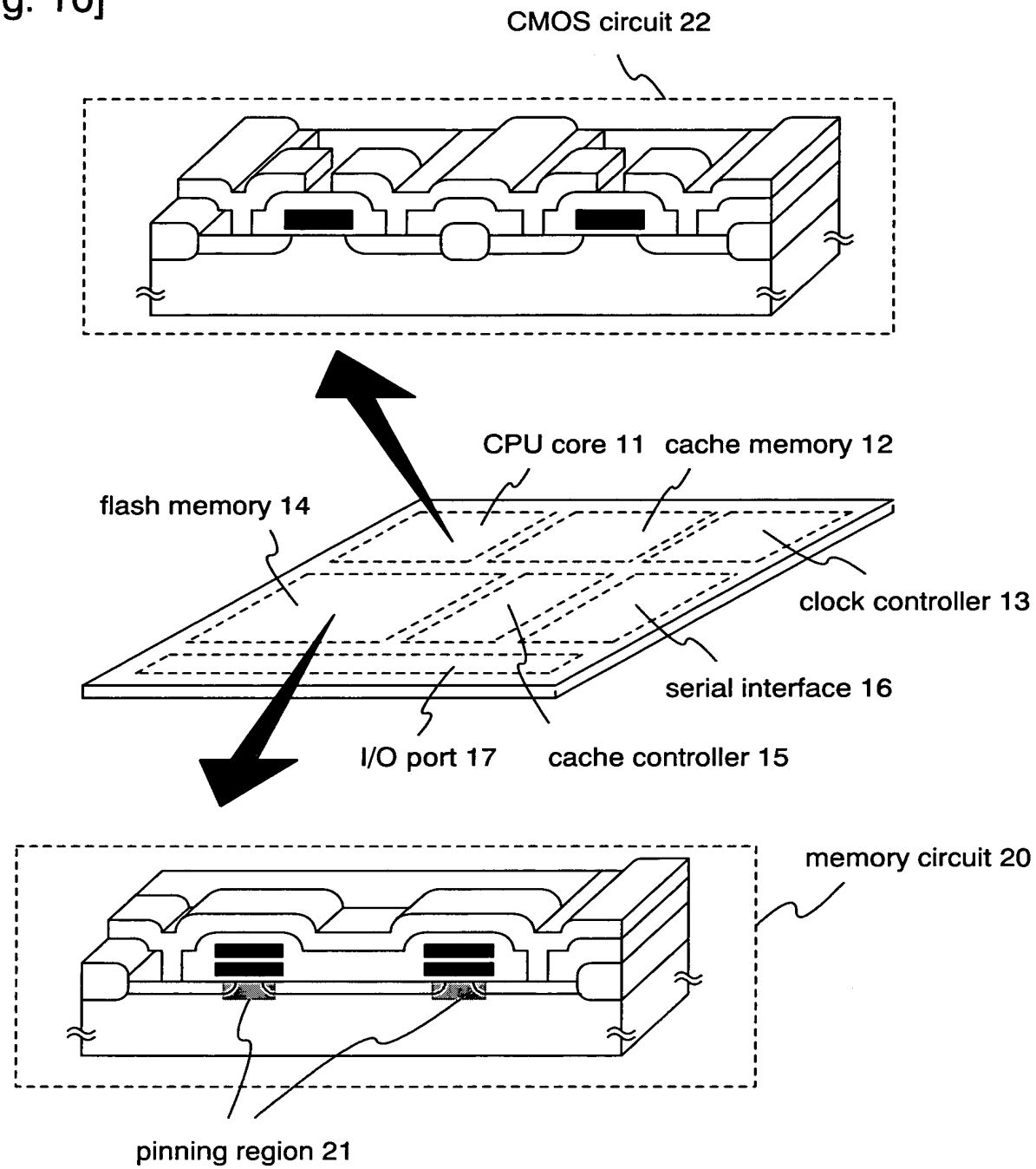


(B)

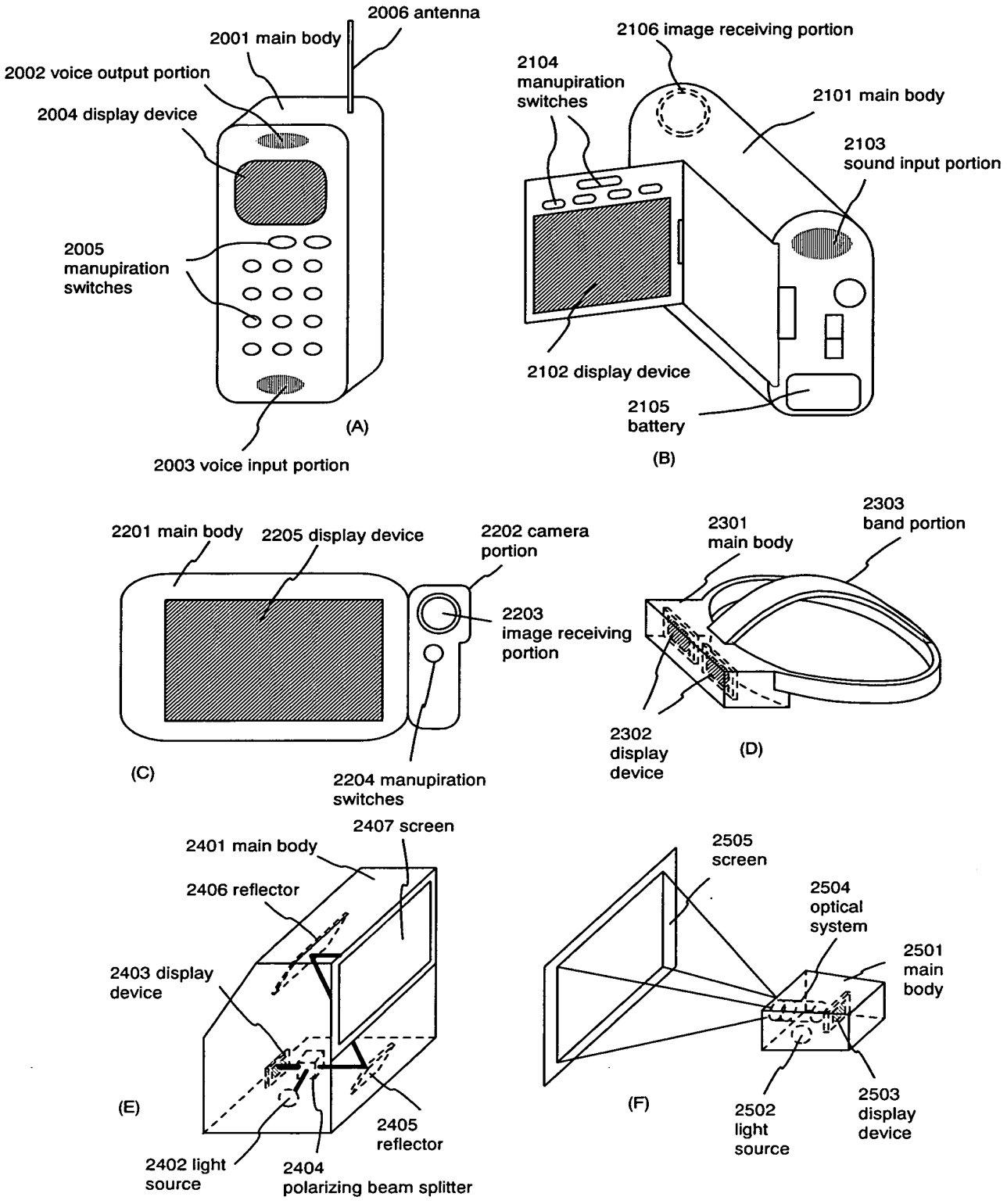
[Reference No.] P003780-01
[Fig. 9]



[Reference No.] P003780-01
[Fig. 10]



[Fig. 11]


2006 antenna
2001 main body
2002 voice output portion
2004 display device
2005 manupiration switches
2003 voice input portion
(A)
2106 image receiving portion
2104 manupiration switches
2101 main body
2103 sound input portion
2102 display device
2105 battery
(B)
2201 main body
2205 display device
2202 camera portion
2203 image receiving portion
2204 manupiration switches
(C)
2303 band portion
2301 main body
2302 display device
(D)
2407 screen
2401 main body
2406 reflector
2403 display device
2405 reflector
2402 light source
2404 polarizing beam splitter
(E)
2505 screen
2504 optical system
2501 main body
2503 display device
2502 light source
(F)